S. T. Rachev, L. Rüschendorf, A transformation property of minimal metrics, *Teor. Veroyatnost. i Primenen.*, 1990, Volume 35, Issue 1, 131–137

ТЕОРИЯ ВЕРОЯТНОСТЕЙ
Том 35 И ЕЕ ПРИМЕНЕНИЯ Выпуск 1
1990



*RACHEV S. T., RÜSCHENDORF L.*

# A TRANSFORMATION PROPERTY OF MINIMAL METRICS

In the first part of this paper we prove an invariance property of minimal metrics with respect to measurable transformations. In the second part this property is used to study the convergence behaviour of various minimal metrics which are related to Kantorovich-type metrics.

Let $(U, \mathcal{A})$, $(V, \mathcal{B})$ be measurable spaces and $\varphi: U \to V$ be a measurable function. Let $\mu: M^1(V \times V) \to [0, \infty]$ be a probability metric on $V$ as defined in [4] (see also [6]), $M^1(V)$ denoting the set of all probability measures on $(V, \mathcal{B})$ then by means of the function $\varphi$ one can define

$$\mu_\varphi: M^1(U \times U) \to [0, \infty] \quad \text{by} \quad \mu_\varphi(Q) := \mu(Q^{(\varphi, \varphi)}), \tag{1}$$

where $Q^{(\varphi,\varphi)}$ is the image of $Q$ under the transformation $(\varphi, \varphi)(x, y) = (\varphi(x), \varphi(y))$. For $P \in M^1(V \times V)$ with marginals $P_1$, $P_2$ let $\hat{\mu}(P) := \inf\{\mu(Q);\ Q \in M^1(V \times V)$ has marginals $P_1, P_2\}$ denote the minimal metric corresponding to $\mu$.

It is easy to see that $\mu_\varphi$ defines a probability metric on $U$. In terms of random variables the above definition can also be written in the following way:

$$\mu_\varphi(X, Y) = \mu(\varphi(X), \varphi(Y)).$$

Recall that $(U, A)$ is called a Borel-space, if there exists an element $B \in \mathfrak{B}^1 = \mathcal{B}(\mathbf{R}^1)$ and a measure isomorphismn $\psi: (U, A) \to (B, B, \mathfrak{B}^1)$. The main aim of this note is to prove the following theorem:

**Theorem.** *Let $(U, A)$ be a Borel-space, $\{v\} \in \mathcal{B}$ for all $v \in V$, then*

$$\hat{\mu}_\varphi(P_1, P_2) = \hat{\mu}(P_1^\varphi, P_2^\varphi) \quad \text{for all} \quad P_1, P_2 \in M^1(U). \tag{2}$$

For the proof of the Theorem we shall need an auxiliary result on the construction of random variables. Let $(M, \mathcal{E}, Q)$ be a probability space, let $S: M \to V$, $Z: M \to [0, 1]$ be independent random variables such that $Q^Z = R(0, 1)$ — the uniform distribution on $[0, 1]$.

**Proposition.** *Let $U$, $V$, $\varphi$ be as in the Theorem and let $P$ be a probability measure on $(U, \mathcal{A})$ such that $P^\varphi = Q^S$. Then there exists a random variable $X: M \to U$ such that*

$$Q^X = P \text{ and } \varphi \circ X = S\ [Q]. \tag{3}$$

P r o o f. Consider at first the special case $(U, \mathcal{A}) = (\mathbf{R}^1, \mathfrak{B}^1)$. Let $\pi: \mathbf{R}^1 \to \mathbf{R}^1$ denote the identity, $\pi(x) = x$, and define $P_s := P^{\pi|\varphi=s}$, $s \in V$; $(P_s)_{s\in V}$ beeing a regular conditional distribution. Let $F_s$ be the right continuous distribution function of $P_s$, $s \in V$. Then

$$F: V \times \mathbf{R}^1 \to [0, 1],\ F(s, x) := F_s(x) \text{ is product-measurable.} \tag{4}$$

For the proof of (4) let $a \in [0, 1]$, then

$$\{(s, x); F_s(x) < a\} = \bigcup_{q\in Q} \{(s, x);\ q \geq x, F_s(q) < a\} =$$
$$= \bigcup_{q\in Q} (\{s \in V; F_s(q) < a\} \times (-\infty, q]),$$

the first equality following from right continuity of $F$ in $x$. Measurability of $F_s(q)$ w.r.t. $s$ implies that the above set is an element of $\mathcal{B} \otimes \mathfrak{B}^1$ and, therefore, the product measurability of $F$.

For $s \in V$ let $F_s^{-1}(x) := \inf\{y; F_s(y) \geqslant x\}$, $x \in (0, 1)$ be the generalized inverse of $F_s$ and define the random variable

$$X := F_s^{-1}(Z). \tag{5}$$

For any $A \in \mathcal{A} = \mathfrak{B}^1$ holds

$$Q^X(A) = \int Q^{X|S=s}(A)\, dQ^S(s)$$

and for the regular conditional distributions we obtain

$$Q^{X|S=s} = Q^{F_S^{-1}(Z)|S=s} = Q^{F_s^{-1}(Z)|S=s} = Q^{F_s^{-1}(Z)}$$

by the independence of $S$ and $Z$. Furthermore, $Q^{F_s^{-1}(Z)} = P_s = P^{\pi|\varphi=s}$ and, therefore, $Q^X(A) = \int P^{\pi|\varphi=s}(A) dP^\varphi(s) = P(A)$. This implies the first relation $Q^X = P$.

The second relation $\varphi \circ X = S\ [Q]$ is implied by

$$Q\{\varphi \circ X = S\} = \int Q^{X|S=s}\{x; \varphi(x) = s\} dQ^S(s) =$$
$$= \int P^{\pi|\varphi=s}\{x; \varphi(x) = s\}\, dP^\varphi(s) = 1$$

since the integrand is equal to one a.s. by the usual properties of regular conditional distributions.

Consider now the general case that $(U, \mathcal{A})$ is a Borel space. Let $\psi: (U, \mathcal{A}) \to (B, B\mathfrak{B}^1)$, $B \in \mathfrak{B}^1$, be a measure isomorphismn and define $P' := P^\psi$, $\varphi' := \varphi \circ \psi^{-1}$. By part one of this proof there exists a random variable $X': M \to B$ such that $Q^{X'} = P'$ and $\varphi' \circ X' = S\ [Q]$. Therefore, with $X := \psi^{-1} \circ X'$, holds $Q^X = P$ and $\varphi \circ X = S\ [Q]$.

R e m a r k. A similar proof holds true under the alternative assumptions:

a) $U$ is a universally measurable separable metric space.

b) $V$ is a separable metric space and for $P_s := P^{\pi|\varphi=s}$, $\pi: U \to U$, $\pi(x) = x$, there exists a productmeasurable process $Y_s := M \to U$, $s \in A$, such that $Q^{Y_S} = P_s$, $s \in V$.

P r o o f o f t h e T h e o r e m. Let $M(P_1, P_2)$ be the set of all probability measures on $U \times U$ with marginals $P_1$, $P_2$. Then

$$\{Q^{(\varphi, \varphi)}; Q \in M(P_1, P_2)\} \subset M(P_1^\varphi, P_2^\varphi)$$

and, therefore,

$$\hat{\mu}_\varphi(P_1, P_2) = \inf\{\mu_\varphi(Q); Q \in M(P_1, P_2)\} = \inf\{\mu(Q^{(\varphi, \varphi)}); Q \in M(P_1, P_2)\} \geqslant$$
$$\geqslant \inf\{\mu(Q); Q \in M(P_1^\varphi, P_2^\varphi)\} = \hat{\mu}(P_1^\varphi, P_2^\varphi).$$

Conversely, let $P \in M(P_1^\varphi, P_2^\varphi)$ let $(M, \mathcal{E}, Q)$ be a measure space with random variables $S, S': M \to V$ such that $Q^{(S, S')} = P$ and rich enough to contain a further random variable $Z: M \to [0, 1]$ uniformly distributed on $[0, 1]$ and independent of $S, S'$. By the Proposition there exist random variables $X, Y: M \to U$ such that $Q^X = P_1$, $Q^Y = P_2$ and $\varphi \circ X = S\ [Q]$, $\varphi \circ Y = S'\ [Q]$. Therefore, $\mu(P) = \mu(\varphi \circ X, \varphi \circ Y) = \mu_\varphi(X, Y)$, implying that

$$\mu_\varphi(P_1, P_2) = \inf\{\mu_\varphi(X, Y); X \sim P_1, Y \sim P_2\} \leqslant$$
$$\leqslant \inf\{\mu(P); P \in M(P_1^\varphi, P_2^\varphi)\} = \mu(P_1^\varphi, P_2^\varphi).$$

Remark. a) The Theorem is known for several special situations (cf. [1—5, 7]). As a typical example for the application of the Theorem let $U = V$ be a Banach space, $d_s(x, y) = | x | x |^{s-1} - y | y |^{s-1} |$, $x, y \in U$ where $s \geqslant 0$ and $x | x |^{s-1} = 0$ for $x = 0$. Let $\mu_s(X, Y) = \mathrm{E} d_s(X, Y)$, $\mu(X, Y) = \mu_1(X, Y)$; then the corresponding minimal metrics $k_s(X, Y) := \hat{\mu}_s(X, Y)$ are called absolute pseudomoments of order $s$ (cf. [3]). The importance of $k_s$ can be explained by the fact that one can obtain upper bounds for Zolotarev's ideal metric $\xi_s$ in terms of $\mathbf{k}_s$ (cf. [3], Theorem 3) and, therefore, obtain rates of convergence in central limit theorems in terms of $\mathbf{k}_s$. By the Theorem (which is trivial in this special case) $\mathbf{k}_s$ can be expressed in terms of the more simple metric $\mathbf{k}_1$, $\mathbf{k}_s(P_1, P_2) = \mathbf{k}_1(P_1^\varphi, P_2^\varphi)$, where $\varphi(x) = x | x |^{s-1}$.

b) An immediate extension of the Theorem is possible to the situation $\varphi, \psi\colon U \to V$ and $\mu_{\varphi,\psi}(X, Y) := \mu(\varphi(X), \psi(Y))$.

Some applications are now discussed in more detail in the following section.

**Applications.** Define the $L_p$-metric in $M^1(V \times V)$

$$L(\mu) := \Big( \int\limits_{V\times V} d^p(x, y(\mu(dx, dy)) \Big)^{1/p}, \; p \geqslant 1,$$

$\mu \in M^1(V \times V)$, assuming that $V$ is a separable metric space with metric $d$. Then, by (1) $L_\varphi$ is a $L_p$-type probability metric in $M^1(U \times U)$ and $\hat{L}_\varphi$ is the corresponding minimal metric. In the next corollary we apply the Theorem in order to get a criterion for $\hat{L}_\varphi$-convergence.

Let $Q, P_1, P_2, \ldots$ be probability measures on $(U, A)$. Denote $\pi_n^\varphi := \pi(P_n^\varphi, Q^\varphi)$, $\pi$ being the Levy — Prohorov metric in $M^1(V)$,

$$D_n^\varphi := D(P_n^\varphi, Q^\varphi) := \Big| \Big( \int\limits_V d^p(x, c) P_n^\varphi(dx) \Big)^{1/p} - \Big( \int\limits_V d^p(x, c) Q^\varphi(dx) \Big)^{1/p} \Big|$$

($c$ is a fixed point in $V$),

$$c(Q^\varphi) := \Big(p \int\limits_V (d(x, c) + 1)^{p-1} Q^\varphi(dx) \Big)^{1/p},$$

$$M(Q^\varphi, N) := \Big( \int\limits_V d^p(x, c) I\{d(x, c) > N\} Q^\varphi(dx) \Big)^{1/p},$$

$$M(Q^\varphi) := \Big( \int\limits_V d^p(x, c) Q^\varphi(dx) \Big)^{1/p}.$$

Let $U$ be a Borel space and $V$ be as above, then we have:

**Corollary 1.** *Let for all* $n \in \mathbf{N}$

$$M(P_n^\varphi) + M(Q^\varphi) < \infty. \tag{6}$$

*Then* $\hat{L}_\varphi(P_n, Q) \to 0$ *as* $n \to \infty$ *if and only if* $P_n^\varphi$ *weakly tends to* $Q^\varphi$ *and* $D_n^\varphi \to 0$ *as* $n \to \infty$. *Moreover, the following quantitative estimates are valid:*

$$\hat{L}_\varphi(P_n, Q) \geqslant \max(D_n^\varphi, (\pi_n^\varphi)^{1+1/p}), \tag{7}$$

$$\hat{L}_\varphi(P_n, Q) \leqslant (1 + 4N)\pi_n^\varphi + 5M(Q^\varphi, N) + (\pi_n^\varphi)^{1/p}(3c(Q^\varphi) + 2^{2+1/p}N) + D_n^\varphi \tag{8}$$

*for each positive* $N$.

Remark. The first part of Corollary 1 follows immediately from relations (7), (8) (for the «if» part put for instance $N = (\pi_n^\varphi)^{-1/(2p)}$). Relations

(7) and (8) establish additionally a quantitative estimate of the convergence of $\widehat{L}_\varphi (P_n, Q)$ to zero.

P r o o f o f (7), (8). Relations (7), (8) follow from (2) and the following inequalities:

$$\widehat{L} (Q_1, Q_2) \geqslant \max (\pi^{1+1/p} (Q_1, Q_2),\ D (Q_1, Q_2)), \tag{9}$$

$$\widehat{L} (Q_1, Q_2) \leqslant (1 + 2N)\, \pi (Q_1, Q_2) + M (Q_1, N) + M (Q_2, N), \tag{10}$$

$$M (Q_1, 2N) \leqslant D (Q_1, Q_2) + 4M (Q_2, N) + \pi^{1+1/p} (Q_1, Q_2) (3c (Q_2) + 2^{1+1/p}N) \tag{11}$$

for each positive $N$ and $Q_1, Q_2 \in M^1 (V)$.

The inequalities (9) and (10) are proved in [1] (for possible extensions see [6]). For the proof of (11) observe that

$$M (Q_1, 2N) \leqslant D (Q_1, Q_2) + \left\{\left|\int_V d^p (x, c)\, I \{d (x, c) \leqslant 2N\} (Q_1 - Q_2) (dx)\right|\right\}^{1/p} + M (Q_2, 2N). \tag{12}$$

In order to estimate the second term, say $I$, in the right-hand side of (12) we denote $f (x) := \min \{d^p (x, c),\ (2N)^p\}$, $g (x) := \min \{2^p d^p (x, 0 (c, N)), (2N)^p\}$, where $O (c, N) := \{x \in U;\ d (x, c) \leqslant N\}$. Then

$$I \leqslant \left|\int_V f (x) (Q_1 - Q_2) (dx)\right|^{1/p} + 2N \left|\int_V I \{d (x, c) > 2N\} (Q_1 - Q_2) (dx)^{1/p} =: I_1 + I_2 \text{ say}. \tag{13}$$

Using the inequality

$$| f (x) - f (y) | \leqslant | d^p (x, c) - d^p (y, c) | \leqslant p \max (d^{p-1} (x, c),\ d^{p-1} (y, c)) \cdot d (x, y), \quad x, y \in V$$

we get for any probability measure $\mu$ on $V^2$ with marginals $Q_1, Q_2$ and $\mu \{d (x, y) > \gamma\} < \gamma$ for some $\gamma \in (0, 1]$:

$$I_1^p = \left|\int_{V^2} (f (x) - f (y))\, \mu (dx, dy)\right| \leqslant \int_{V^2} | f (x) - f (y) |\, I \{d (x, y) \leqslant \gamma\} \mu (dx, dy) + \int_{V^2} (| f (x) + f (y) |)\, I \{d (x, y) \geqslant \gamma\}\, \mu (dx, dy) \leqslant \gamma c (Q_2)^p + 2 (2N)^p \gamma.$$

Let $K := K (\mu) := \inf \{\gamma > 0;\ \mu (d (x, y) > \gamma) < \gamma\}$ denote the Ky Fan distance in $M^1 (V \times V)$ then

$$I_1 \leqslant K^{1/p} [c (Q_2)^p + 2 (2N)^p]^{1/p} \leqslant K^{1/p} [c (Q_2) + 2^{1+1/p} N]. \tag{14}$$

Furthermore,

$$I_2 = \left|\int_V (2N)^p I \{d (x, c) > 2N\} (Q_1 - Q_2) (dx)\right|^{1/p} \leqslant \left[\int_V (2N)^p I \{d (x, c) > 2N\}\, Q_1 (dx)\right]^{1/p} + M (Q_2, 2N).$$

If $d (x, c) > 2N$, then $d (x, O (c, N)) \geqslant N$, and, therefore,

$$\left[\int_V (2N)^p I \{d (x, c) > 2N\}\, Q_1 (dx)\right]^{1/p} \leqslant \left[\int_V g (x)\, Q_1 (dx)\right]^{1/p} \leqslant \left[\left|\int_V g (x) (Q_1 - Q_2) (dx)\right|\right]^{1/p} + \left[\int_V g (x)\, Q_2 (dx)^{1/p} = I_{21} + I_{22}.$$

We obtain

$$I_{22} \leqslant \left[\int_V (2N)^p I\{d(x,c) > N\} Q_2(dx)\right]^{1/p} \leqslant$$

$$\leqslant \Big( \int_{\{d(x,c)>N\}} 2^p d^p(x,c) Q_2(dx)\Big)^{1/p} = 2M(Q_2, N)$$

and hence the inequality

$$|g(x) - g(y)| \leqslant 2^p \,|d^p(x, O(c, N)) - d^p(y, O(c, N))| \leqslant$$
$$\leqslant 2^p p \max [d^{p-1}(x, O(c, N)), d^{p-1}(y, O(c, N))]\, d(x, y)$$

implies

$$I_{21} \leqslant \Big[\int_{V^2} |g(x) - g(y)| I\{d(x,y) \leqslant \gamma\} \mu(dx, dy)\Big]^{1/p} +$$

$$+ \Big[\int_{V^2} (|g(x)| + |g(y)|) I\{d(x,y) > \gamma\} \mu(dx, dy)\Big]^{1/p} \leqslant 2\gamma^{1/p} c(Q_2) + 2^{1+1/p} N\gamma^{1/p}.$$

Alltogether, we get (as for the bounds of $I_1$)

$$I_2 \leqslant 3M(O_2, N) + 2K^{1/p} c(Q_2) + 2^{1+1/p} N K^{1/p} \tag{15}$$

for any $K = K(\mu)$. By the Strassen theorem $\pi = \hat{K}$ and hence $K$ can be replaced in the bounds by $\pi$ implying

$$I \leqslant 3M(Q_2, N) + \pi^{1/p}(3c(Q_2) + 2^{2+1/p} N). \tag{16}$$

This implies relation (11).

We can extend Corollary 1 considering the compound probability distance

$$\mathscr{H}(\mu) := \int_{V^2} H(d(x, y))\, \mu(dx, dy), \quad \mu \in M^1(V \times V), \tag{17}$$

where $H(t)$ is a nondecreasing continuous function on $[0, \infty)$ vanishing at zero (and only there) and satisfying the Orlicz-condition

$$\sup\{H(2t)/H(t);\ t > 0\} < \infty \quad \text{(see [6]).} \tag{18}$$

**Corollary 2.** *Assume that* $\int H(d(x, y))(P_n^\varphi + Q^\varphi)(dx) < \infty$. *Then the convergence* $\hat{\mathscr{H}}_\varphi(P_n, Q) \to 0$ *as* $n \to \infty$ *is equivalent to the following relations*:

a) $P_n^\varphi$ *tends weakly to* $Q^\varphi$ *as* $n \to \infty$ *and*

b) $\lim\limits_{N\to\infty} \overline{\lim\limits_{n}} \int_V H(d(x, c)) I\{d(x, c) > N\} P_n^\varphi(dx) = 0.$

Note that the Orlicz-condition (18) implies a power growth of the function $H$. In order to consider functions $H$ in (17) with exponential growth we introduce the class SB of «subbounded» rv's $\xi$. Define

$$\xi \in SB \Leftrightarrow \tau(\xi) := \inf\{a > 0;\ \mathbf{E}\exp(\lambda\xi) \leqslant \exp \lambda a \text{ for all } \lambda > 0\} =$$
$$= \sup_{\lambda>0} \frac{1}{\lambda} \ln \mathbf{E}\exp(\lambda\xi) < \infty. \tag{19}$$

Obviously all bounded rv's belong to SB. By the Hölder inequality one gets

$$\tau(\xi + \eta) \leqslant \tau(\xi) + \tau(\eta) \tag{20}$$

and hence if $\mu \in M^1(V \times V)$, and $(Y_1, Y_2)$ is a pair of $V$-valued rv's with joint distribution $\mu$, then

$$\tau(\mu) := \tau(d(Y_1, Y_2)) \tag{21}$$

determines a compound probability metric on $M^1(V \times V)$ (see [1, 4, 6]). The next corollary of the Theorem gives us a criterion for $\hat{\tau}_\varphi$-convergence.

**Corollary 3.** *Let $X_n$, $n = 1, 2, \ldots$, and $Y$ be valued rv's with distributions $P_n$ and $Q$ respectively and let $\tau(d(\varphi(X_n), c)) + \tau(d(\varphi(Y), c)) < \infty$. Then the convergence $\hat{\tau}_\varphi(P_n, Q) \to 0$ as $n \to \infty$ is equivalent to the following relations:*

a) *$P_n^\varphi$ tends weakly to $Q^\varphi$ and*

b) $\lim\limits_{N\to\infty} \overline{\lim\limits_{n}}\, \tau(d(\varphi(X_n), c) I\{d(\varphi(X_n), c) > N) = 0$.

P r o o f. As in Corollary 1 the assertion of Corollary 3 is a consequence of Theorem 1 and the following inequalities (22) — (24) valid for any $V$-valued random variables $Y_1$ and $Y_2$ with distributions $Q_1$ and $Q_2$ respectively:

$$\pi^2(Q_1, Q_2) \leqslant \hat{\tau}(Q_1, Q_2); \tag{22}$$

$$\tau(d(Y_1, c) I\{d(Y_1, c) > N\}) \leqslant 2\hat{\tau}(Q_1, Q_2) + 2\tau(d(Y_2, c) I\{d(Y_2, c) > N/2\}), \tag{23}$$

$$\hat{\tau}(Q_1, Q_2) \leqslant \pi(Q_1, Q_2)(1 + 2N) + \tau(d(Y_1, c) I\{d(Y_1, c) > N\}) + \tau(d(Y_2, c) I\{d(Y_2, c) > N\}) \text{ for all } N > 0,\ c \in V. \tag{24}$$

P r o o f o f (22). By the Strassen theorem it is enough to prove that $\tau(\mu) \geqslant K^2(\mu)$ for $\mu \in M^1(V \times V)$ with marginals $Q_1$, $Q_2$. Let $\xi = d(Y_1, Y_2)$, where $(Y_1, Y_2)$ has the distribution $\mu$, and $\tau(\xi) < \varepsilon^2 \leqslant 1$, then

$$\mathbf{P}(\xi > \varepsilon) \leqslant \frac{\mathbf{E}e^\xi - 1}{e^\varepsilon - 1} \leqslant \frac{e^{\tau(\xi)} - 1}{e^\varepsilon - 1} \leqslant \frac{e^{\varepsilon^2} - 1}{e^\varepsilon - 1}.$$

Letting $\varepsilon^2 \to \tau(\xi)$ we get (22).

P r o o f o f (23). Note that the inequality $\xi \leqslant \eta$ with probability one implies $\tau(\xi) \leqslant \tau(\eta)$. Hence

$$\begin{aligned} &\tau(d(Y_1, c) I\{d(Y_1, c) > N\}) \leqslant \\ &\leqslant \tau[(d(Y_1, Y_2) + d(Y_2, c)) I\{d(Y_2, c) + d(Y_1, Y_2) > N\}] \leqslant \\ &\leqslant \tau\left[(d(Y_1, Y_2) + d(Y_2, c)) \max\left(I\left\{d(Y_2, c) > \frac{N}{2}\right\}, I\left\{d\left(Y_1, Y_2\right) > \frac{N}{2}\right\}\right)\right] \leqslant \\ &\leqslant 2\tau\left(d(Y_1, Y_2) I\left\{d(Y_1, Y_2) > \frac{N}{2}\right\}\right) + 2\tau\left(d(Y_2, c) I\left\{d(Y_2, c) > \frac{N}{2}\right\}\right) \leqslant \\ &\leqslant 2\tau(d(Y_1, Y_2)) + 2\tau\left(d(Y_2, c) I\left\{d(Y_2, c) > \frac{N}{2}\right\}\right). \end{aligned}$$

Passing to the minimal metrics we get (23).

P r o o f o f. (24) For each $\delta$ holds

$$\tau(d(Y_1, Y_2)) \leqslant \tau(d(Y_1, Y_2) I\{d(Y_1, Y_2) \leqslant \delta\}) + \tau(d(Y_1, Y_2) I\{d(Y_1, Y_2) > \delta\}) =: I_1 + I_2, \text{ say.}$$

For $I_1$ we get the estimate

$$I_1 = \sup_{\lambda > 0} \frac{1}{\lambda} \ln \mathbf{E} \exp(\lambda d(Y_1, Y_2) I\{d(Y_1, Y_2) \leqslant \delta\}) \leqslant \sup_{\lambda > 0} \frac{1}{\lambda} \ln \mathbf{E} \exp \lambda\delta = \delta.$$

For $I_2$ we get:

$$\begin{aligned} I_2 &\leqslant \tau([d(Y_1, c) + d(Y_2, c)\} I\{d(Y_1, Y_2) > \delta\}) \leqslant \\ &\leqslant \tau(d(Y_1, c) I\{d(Y_1, Y_2) > \delta\}) + \\ &+ \tau(d(Y_2, c) I\{d(Y_1, Y_2) > \delta\}) =: A_1 + A_2. \end{aligned}$$

Furthermore,

$$\begin{aligned} A_1 &\leqslant \tau(d(Y_1, c) I\{d(Y_1, Y_2) > \delta\} I\{d(Y_1, c) \leqslant N\}) + \\ &+ \tau(d(Y_1, c) I\{d(Y_1, Y_2) > \delta\} I\{d(Y_1, c) > N\}) \leqslant \\ &\leqslant NP(d(Y_1, Y_2) > \delta) + \tau(d(Y_1, c) I\{d(Y_1, c) > N\}). \end{aligned}$$

Hence if $K(Y_1, Y_2) < \delta$, then

$$\tau(d(Y_1, Y_2)) \leqslant (1 + 2N)\delta + \tau(d(Y_1, c) I\{d(Y_1, c) > N\}) + \\ + \tau(d(Y_2, c) I\{d(Y_2, c) > N\}).$$

Letting $\delta \to K(Y_1, Y_2)$ and passing to the minimal metrics we obtain (24).

For another possible line of applications of the Theorem cf. [7].

## REFERENCES

1. *Золотарев В. М.* Метрические расстояния в пространствах случайных величин и их распределений.— Матем сб., 1976, т. 101 (143), № 3, с. 416—454.
2. *Золотарев В. М.* Аппроксимация распределений сумм независимых случайных величин со значениями из бесконечномерных пространств.— Теория вероятн. и ее примен., 1976, т. XXI, в. 4, с. 741—758.
3. *Золотарев В. М.* О псевдомоментах.— Теория вероятн. и ее примен., 1979, т. XXIII, в. 2, с. 284—294.
4. *Золотарев В. М.* Вероятностные метрики.— Теория вероятн. и ее примен., 1983, т. XXVIII, в. 2, с. 264—287.
5. *Игнатов Ц., Рачев С. Т.* Минимальность идеальных вероятностных метрик.— В сб.: Проблемы устойчивости стохастических моделей. М.: ВНИИСИ, 1983, с. 36—48.
6. *Rachev S.* Extreme functionals in the space of probability measures.— Lect. Notes Math., 1985, B. 1155, S. 320—348.
7. *Rüschendorf L.* The Wasserstein distance and approximation theorems.— Z. Wahrscheinlichkeitstheor. verw. Geb., 1985, B. 70, H. 1, S. 117—129.

Поступила в редакцию
16.VI.1986